Two Copy Testing Techniques: The Cloze Procedure and the Cognitive Complexity Test

George M. Zinkhan, *University of Houston*
Claude R. Martin, Jr., *The University of Michigan*

The theory behind two copy testing techniques—the cloze procedure and the cognitive complexity text—is reviewed; and three hypotheses are developed for testing. These two techniques work especially well when used together as they provide a way to match message properties with the audience's abilities. In a discriminant analysis setting, support is found for all three hypotheses, and a substantial portion of the variance in the criterion variable—advertising recall—is taken into account. The results of predictions to a validation sample are also encouraging.

Introduction

Advertisers are interested in designing effective promotional campaigns; but they are also interested in the effectiveness of individual ads. Multiple methods are required for the task of testing advertising copy, and there is always a need for the refinement and improvement of testing techniques [12]. The purpose of this paper is to introduce two copy testing techniques which work well in tandem—the cognitive complexity test and the cloze procedure. The former provides a measure of audience characteristics, while the latter provides a measure of advertising structure. Conceptual issues related to copy testing are discussed; hypotheses are developed and empirically tested with respect to print advertisement for calculators.

Structure of the Ad

In order to achieve its objectives, an advertisement must have a well-defined design or structure [7]. Anderson and Jolson [3] studied the effect that the use of technical or complex words has on advertising recall. These two researchers propose that the more the wording of the ad approaches the language of the particular sector to which it is directed, the easier it becomes for the factors of comprehension and interest to

retain the readers' attention. Thus Anderson and Jolson investigate the effects of individual complex words in an advertisement. The study planned here goes one step further in that the complexity or structure of entire sentences or paragraphs is examined. The way in which words, sentences, and paragraphs fit together to form identifiable patterns is investigated. To rephrase Anderson and Jolson, the more the structure of the ad is comprehensible or apparent to the particular sector to which it is directed, the easier it becomes for the ad to retain the reader's attention. This idea of "comprehending structure" can be examined in more detail through the application of a particular theory of consumer aesthetics— the theory of requiredness.

Requiredness applies to aesthetic objects which involve a sequential string of individual components (e.g., music, literature, and dance). These are objects which cannot be perceived all at once in a Gestalt-type manner, but instead reveal their segments only one at a time. Print advertisements can be viewed as a sequential string of words and can also be viewed as aesthetic objects. Aesthetic objects which possess this sequential string ordering may have their form partially predetermined by the style in which the creator has chosen to work [17]. After a sequence of words has occurred, the next word is, to a certain degree, predetermined or "required" by the pattern or structure of preceding words. To create a successful artwork, then, the writer must set up a certain structure. This component of requiredness may be what distinguishes successful or good pieces of writing from writing of lesser quality.

Measuring Advertising Structure

The cloze test measures an individual's ability to provide deleted words in a passage of writing. As such, the cloze procedure can be viewed as a measure of verbal uncertainty [8]. The cloze test is similar to other measures of readability (such as Flesch test, Gunning's Fog Index, or the Dale–Chall formula); but it is superior in that cloze scores are the result of a reader interacting with a written passage, rather than representing a score based on the length of sentences or the length of words within sentences.

The cloze procedure measures the likeness between the patterns a writer has used and the patterns the reader is anticipating while reading. The cloze method does not deal directly with specific meaning. Instead, it repeatedly samples the extent of likeness between the language patterns used by the writer to express what he meant and those possible different patterns which represent readers' guesses at what they *think* the writer meant [15]. The cloze procedure seems particularly well suited to

operationalize the theory of requiredness and the readers' "comprehension of advertising structure," as outlined above.

The cloze procedure can be administered as follows. After a random start, every sixth word in the passage is deleted. The "mutilated" passages are then reproduced with all missing words replaced by standard-size blanks. Subjects are asked to "close up the gaps" in the passages by guessing the identities of the missing words and writing their guesses in the corresponding blanks. Each time a subject correctly guesses a missing word he scores one point; an individual's "cloze score" for any particular passage is simply the total number of words guessed correctly [16]. MacGinitie [9] and Alderson [2] have found that the amount of context between cloze gaps does not have any significant effect on the predictability of the deleted word, providing that at least five words of context are available. In other words, it is best to delete every sixth word.

One use of the cloze test is to measure readability [16]. The more that a reader is "in tune" with the verbal patterns of the writer, the better the reader is able to understand a particular passage. High levels of understanding should be related to high levels of advertising recall.

H_1: The better an individual is able to predict the verbal structure of an advertisement, the more the individual will be able to recall information from that ad.

Cognitive Complexity

Cognitive complexity refers to the number and sophistication of cognitive *structures* that a given individual may possess. A distinction is made here between cognitive content and cognitive structure. The content of cognition consists of concepts of objects and their attributes [14]. The structure of cognition refers to the relationships among these objects. Cognitive complexity, therefore, is based not on the content of cognition but rather on the presence of structures for organizing this content.

Research indicates that cognitive complexity is domain specific. For example, a person might be cognitively complex with respect to calculators and yet, this same individual may be cognitively simple with respect to automobiles. Those with highly differentiated cognitive structures should be able to keep associations, messages, and sources of information separate within their minds [18]. In other words, forgetting should be less of a problem for those with highly differentiated cognitive structures.

H_2:   The higher the individual's level of cognitive complexity, the more that individual will be able to recall the advertising message.

Measuring Cognitive Complexity

Cognitive complexity is operationalized in a manner outlined by Scott [14] using two stages. First, in a free format, respondents list all of the important attributes that they can think of associated with a product. Second, respondents put these attributes into groups that are similar. A respondent can make as many or as few groups as seem appropriate. A measure of cognitive complexity can be obtained according to a formula derived from information theory

$$R = \left(\log_2 (n) - \frac{1}{n} \sum_{i=1}^{n} n_i \log_2 (n_i) \right) [\log_2 (n)]^{-1},$$

where n is the total number of attributes and n_i is the number that appears in a particular combination of groups. A subject's R score is used as a measure of cognitive complexity in this investigation; see Zinkhan and Martin [19] for more details concerning the calculation and relevance of this measure.

The Advertisements

The stimulus object consists of two 250-word long print advertisements for fictitious calculator brands—the Computron R-55 and the El-tronic 38. (The ads were written by professional copywriters at Hanish Associates in Florida. The layout work was done by the advertising department at Jacobson's in Jackson, Michigan; the production work was done by Winkleman's advertising department in Detroit, Michigan.) Since fictitious brands were used, this constitutes a new product situation. The actual product information contained in the ads is determined on the basis of two pretests; and the two advertised products are designed to be similar to one another in terms of attributes. The two calculator ads are also similar in terms of sentence length and word difficulty. The two ads are designed to differ from one another in terms of structure—one ad being complex in structure and the second ad being more simple in structure.

   In order to ensure that the two ads are comparable in quality, a panel of four advertising experts rated the ads on multiple-item scales to assess product positioning and copy execution. The two ads do not significantly differ along these dimensions.

When administering the cloze procedure for the two ads, every sixth word is deleted so that there are in total 43 blanks which the respondent attempts to "close up." The Kuder–Richardson Formula 20 (KR-20) can be used to assess the reliability of the cloze procedure. Coefficient $\propto$ for the more simply written ad is 0.76, while coefficient $\propto$ for the more complex ad is 0.71. The $\propto$ of 0.71 is the lowest attained in this study, but it is still well above the lower limit of 0.50 recommended by Nunnally [11].

Advertising Response

When assessing the effectiveness of an advertisement, it is necessary to consider both product-related and ad-related measures [6]. One such ad-related measure is immediate or initial response to the advertisement. Another term for this phenomenon is aesthetic response. In other words, does the reader respond favorably to the ad on first exposure?

It is possible for affect to lead to cognition. Pleasurable feelings about the ad can lead to favorable thoughts about the product advertised. Of course unenjoyable ads can also be well remembered. It makes sense, though, that a pleasurable ad should provide positive reinforcement for the product advertised.

H_3 The more a person enjoys a particular advertisement, the better that person will be able to recall information about that particular ad.

Measuring Advertising Response

Aesthetic response to advertising focuses on the emotional component of communication effects. Unlike measurements of learning and comprehension, it does not directly concern itself with the retention of claims, slogans, or other factual material. Aesthetic response has to do with how people feel after seeing an ad rather than what they know.

Immediately after reading an advertisement, respondents in our study indicate how enjoyable the ad is to read. Four items are used to tap this notion of advertising response: enjoyment, likability, persuasiveness, and interestingness. An eight-point Likert-type scale is used, and these items are summed to form a composite index. Thus, one enjoyment score is obtained for each of the two calculator ads investigated. Estimates for coefficient $\propto$ associated with these indices—0.91 and 0.90—are quite substantial.

At first glance, it may seem somewhat strange that persuasiveness is included with these other measures. Bartos [4] provides a rationale for this; and the reliability estimates obtained here indicate that persuasive-

ness is an appropriate item to include in the composite which we call aesthetic response.

Measuring Advertising Effectiveness

The purpose of a copy testing technique is to assess the effectiveness of an advertisement. The most commonly used measure of advertising effectiveness is employed here: advertising recall.

The number of facts that a respondent can recall about an advertised product is one possible measure of ad recall. Claycamp and Liddy [5] use this measure to develop their Ayer Model of New Products. To operationalize this concept in an aided fashion, respondents answer true/false questions relating to factual material contained in the ad. Recall measures are taken 1 day after ad exposure. A composite index is formed by summing up the number of correct responses; the KR-20 formula can be used to assess reliability. In our study coefficient $\propto$ associated with the simple ad was 0.81, and coefficient $\propto$ associated with the more complex ad was 0.85. These estimates $\propto$ alpha are sizable and acceptable, according to Nunnally's [11] guidelines.

Sample

Two hundred thirty-eight students at a major midwestern university were selected to take part in the main survey. All of the subjects were paid $10.00 for their participation. Students seem to be especially appropriate for the purposes of the present study since they are interested in and are users of calculators as a product class.

Procedure

Questions were administered in four different sessions. Cognitive complexity was measured in the first session; the cloze procedure was administered during the second session. In the third session, respondents were exposed to four advertisements—two of which were for calculators. After exposure to each of the four ads, the respondents answered questions about the ads designed to assess ad response. During the fourth and last session, the measures of advertising recall were taken. The fourth session occurred one day after the third. The time interval between the other sessions was 1 week in duration.

Since there was a chance that the data gathering instruments would affect one another, several checks were performed. In order to test for ordering effects, half of the respondents were exposed to the Computron ad first and half were exposed to the El-tronic ad first. One-way analysis of variance—performed on ad recall scores—indicates no significant ($p <$

0.05) ordering effect for either the Computron recall scores ($F = 0.593$: *df*. 1, 138) or the El-tronic recall scores ($F = 0.851$; *df*. 1, 138).

In order to determine if exposure to the cloze procedure could increase recall scores, a second group of 115 subjects was exposed to the ads and tested for recall without taking part in a cloze procedure. Analysis of variance results show no significant ($p < 0.05$) difference between these two groups in terms of Computron recall scores ($F = 1.229$; *df*. 1, 253) and in terms of El-tronic recall scores ($F = 1.433$; *df*. 1, 253).

Results

Advertisers are interested in discriminating between an effective message and a less effective message. The objective of a discriminant analysis is to classify people (or objects), by a set of independent variables, into one or more mutually exclusive and exhaustive categories. In this case the dependent variable, advertising recall, is dichotomized as follows:

> 0, subject recalls the simple ad best;
> 1, subject recalls the complex ad best.

The candidate independent variables include: cognitive complexity scores, cloze procedure scores, and ad response scores. The results of this discriminant analysis are presented in Table 1. An *F* test indicates that the assumption of equal covariance matrices is not significantly violated. The Mahalanobis D^2 is sizable (0.57) and significant beyond the 0.001 level.

The ad recall variable is coded so that a success (1) is equivalent to superior recall for the El-tronic 38 advertisement; a failure (0) is equivalent to superior recall for the Computron R-55 advertisement. The weights and signs of the discriminant function are again consistent with the hypotheses previously developed.

Afifi and Azen [1] and Rao [13] outline a way to test the significance of discriminant variables by using a stepwise procedure. At each successive step, the conditional distribution of each variable not entered given the variables entered is considered. The next variable to enter is the one for which the mean values of the conditional distribution in the two populations are ''most different.'' This difference can be measured by a one-way analysis of variance *F* statistic. The *F* statistics for the four independent variables—ad response (Computron), ad response (El-tronic), cognitive complexity, and cloze (El-

Table 1: Discriminant Function: Ad Recall as the
Dependent Variable[a]

Equality of Covariances: $df = 10, 65,066; F = 1.866$[A]

Predictor Variables	Weight	F-Stat	Means Recall Computron Best	Recall El-tronic Best
Constant	−1.164			
Ad Response (Computron)	−0.602	10.31[A]	18.23	12.75
Ad Response (El-tronic)	0.756	6.58[A]	14.18	17.57
Cognitive Complexity	0.407	4.18[A]	0.409	0.714
Cloze (El-tronic)	0.326	4.67[A]	16.77	20.03

$N = 140; D^2 = 0.570; F\text{-Stat} = 4.64$[B]

[A]$p < 0.05$; [B]$p < 0.001$

Ad Recall Validation Sample
Predicted Group Membership[a]

A G M		Number of cases	Recall computron better	Recall El-tronic better
C R E				
T O M	Recall computron better	56	61.1%	38.9%
U U B				
A P E				
L R				
S	Recall El-tronic better	42	45.2%	54.8%
H				
I				
P				

[a] Proportional chance criterion = 51.3%. Total percent correctly classified = 58.2%.

tronic)—are displayed in Table 1. Since all four univariate F statistics
are significant beyond the 0.05 level, support is found for all three
hypotheses.

The person who enjoys the Computron ad a great deal is able to recall
the Computron ad better than the El-tronic ad. Similarly, the person
who enjoys the El-tronic ad quite a bit tends to recall the El-tronic ad
better than the Computron ad. The person who is cognitively complex
recalls the more complex ad better than the simpler ad. And finally, the
person who successfully predicts the structure of a complex ad recalls

the complex ad better than the simple ad. In sum, support is established for the three hypotheses.

A holdout sample of 98 persons is used to assess predictive validity. Table 1 presents the classification results. A total of 58.2% are correctly classified in terms of their relative amounts of recall for the Computron vs. the El-tronic advertisements. This is an improvement over the expected proportional chance criterion of 51.3%.

Discussion

The long-range goals of advertising research are of two types. The first is the development of an adequate theory of how advertising influences consumers. The second is the development of reliable and valid research techniques to evaluate advertising effectiveness and to test theoretical formulations. This research has made progress toward both of these goals.

The theory development related to cognitive processes and consumer aesthetics is supported by the results. Support is found for all three hypotheses, and a substantial portion of the variance in the criterion variable is taken into account. The results of the discriminant analysis give a definite boost to the copy testing techniques investigated here. The cloze procedure, as applied to print ads, supplies a technique to evaluate competing ads, both in terms of overall success and in terms of appropriateness to the target audience. Additionally, it is important to note that the cloze procedure has been used for pictures [10]; therefore future researchers could apply this procedure to assess the graphic portions of advertisements.

The cognitive complexity test, elaborated upon here, is a useful companion to the cloze test. When used in conjunction these two tests can provide a manager with (i) an idea concerning the nature of the receptivity and sophistication of the target and audience, and (ii) an idea of the best way to appeal to this target audience, both in terms of advertising structure and in terms of proper language to use. The isolation of target audiences on the basis of cognitive complexity may help to design better copy and media strategies aimed at communication.

A widespread apothegm regarding print advertising copy is "keep it simple." The assumed premise is that many readers have such limited vocabularies or low thresholds of boredom that the copywriter must use readily understood words [3]. However, the results of this study indicate that it is not only the words which must be kept simple. The structure of the sentences must also be kept simple. It makes sense that a highly repetitive or predictable ad would be more readily assimilated

by the entire target audience than would a less repetitive or predictable ad. This is where the cloze test fits in. The cloze test provides a measure of the complexity, repetitiveness, or predictability of the advertisement. Cognitive complexity, combined with the cloze test, conveys to the manager some idea as to who would be bored by an ad that is too simple or repetitive. In other words, the individual abilities of an ad reader to "appreciate" advertising copy or structure can be taken into account.

The cloze procedure and the measure of repetition that the cloze procedure provides could lead managers to an understanding of the best way to present information to consumers. The concept of cognitive complexity could sensitize managers to the differing information needs that different segments of the population may possess. Thus, these two techniques provide a way to match the characteristics and abilities of the target audience with the methods and measures used to assess advertising copy. Both the cloze test and the cognitive complexity test are easy to administer and are particularly relevant to the investigation of often elusive marketing phenomena.

References

1.   Afifi, A. A., and Azen, S. P., *Statistical Analysis: A Computer Oriented Approach*. Academic Press, New York, 1972.

2.   Alderson, J. Charles, The Effect on the Cloze Test of Changes in Deletion Frequency, *J. Res. Reading* 2(2) (1979): 108-19.

3.   Anderson, Rolph E. and Jolson, Marvin A., Technical Wording in Advertising: Implications for Market Segmentation, *J. Marketing* 44 (Winter 1980): 57-66.

4.   Bartos, Rena, Ads that Irritate May Erode Trust in Advertised Brands, *Harv. Bus. Rev.* 59 (May-June 1981): 138-140.

5.   Claycamp, Henry J. and Liddy, Lucien E., Prediction of New Product Performance: An Analytical Approach, *J. Marketing Res.* 6 (November 1969): 414-20.

6.   Goodwin, Stephen, and Etgar, Michael, An Experimental Investigation of Comparative Advertising: Impact of Message Appeal, Information Load, and Utility of Product Class, *J. Marketing Res.* 17 (May 1980): 187-202.

7.   Hefzallah, Ibrahim M., and Maloney, W. Paul, Are There Only Six Kinds of TV Commercials?, *J. Advertis. Res.* 19 (August 1979): 57-62.

8.   Holbrook, Morris B., Effect of Subjective Interletter Similarity, Perceived Word Similarity, and Contextual Variables on the Recognition of Letter Substitutions in a Proofreading Task, *Percept. Motor Skills* 47 (1978): 243-58.

9.   MacGinitie, W. H., Contextual Constraint in English Prose, Unpublished Ph.D. thesis: Columbia University, 1960.

10. Nicki, R. M. and Moss, Virginia, Preference for Non-Representational Art as a Function of Various Measures of Complexity, *Can. J. Psychol.* 29(3) (1975): 237-49.

11. Nunnally, Jun C., *Psychometric Theory*. McGraw-Hill, New York, 1978.

12. Plummer, Joseph T., Copy Research in the Coming Decade, in (Morris B. Holbrook and Elizabeth Hirschman, eds.) *Consumer Esthetics and Symbolic Consumption*. Association for Consumer Research, Ann Arbor, 1980, pp. 86–89.

13. Rao, C. R., *Advanced Statistical Methods in Biometric Research*. Wiley, Inc., New York, 1952.

14. Scott, William A., Cognitive Complexity and Cognitive Flexibility, *Sociometry* 25 (1962): 405–14.

15. Taylor, Wilson L., "Cloze Procedure": A New Tool for Measuring Readability, *Journalism Quart.* 30 (Fall 1953): 415–33.

16. Taylor, Wilson L., Recent Developments in the Use of "Cloze Procedure," *Journalism Quart.* 33 (Winter 1956): 42–48, 99.

17. Wallendorf, Melanie, Zinkhan, George, and Zinkhan, Lydia S., Cognitive Complexity and Aesthetic Preference, in (Morris B. Holbrook and Elizabeth Hirschman, eds.) *Consumer Esthetics and Symbolic Consumption*. Association for Consumer Research, Ann Arbor, 1980, pp. 52–59.

18. Zinkhan, George M., The Influence of Cognitive Complexity and Copy Structure on Advertising Effectiveness, Unpublished Doctoral Dissertation: The University of Michigan, 1981.

19. Zinkhan, George M. and Martin, Claude R., The Attitudinal Implications of a New Brand's Name in (A. Mitchell, ed.) *Advances in Consumer Research*. Association for Consumer Research, Ann Arbor, 1982, Vol. IX.